Inter and Intra-Rater Reliability of Soft Tissue Palpation Scoring in the Equine Thoracic Epaxial Region.

Megan Merrifield-Jones[a], Gillian Tabor[a*] and Jane Williams[a]

[a]Hartpury University, Gloucester, UK, GL19 3BE

**Corresponding author* gillian.tabor@hartpury.ac.uk

Ethics: Ethical approval gained from the Hartpury University Ethics Committee
Conflicts of interest: none
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Background: Back pain is a significant factor for horses and is challenging for professionals to diagnose, with assessment frequently utilising subjective tools such as manual palpation. Reliable and valid objective measures are required and pressure algometry (PA) has been investigated as an assessment tool, however it has limitations and other more realistic methods may be better suited for the task.

Objectives: To establish inter- and intra-rater reliability for PA, Flexiforce Sensor (FFS) and Manual Palpation for equine epaxial soft tissue, measuring mechanical nociception threshold responses.

Methods: In group 1, ten horses underwent three repeated tests with PA and FFS, and once for manual palpation, with three ACPAT Chartered Physiotherapists in the right thoracic epaxial region. Group 2 followed the same protocol using one ACPAT Chartered Physiotherapist and 22 horses. The order of palpation was randomly applied for each test and each experimenter.

Results: Manual palpation showed excellent inter-rater reliability with no significant differences between scores (p=0.64; ICC 90.0%). PA (p=0.002) and FFS (p=0.025) scores significantly differed between experimenters. Intra-rater testing showed significant differences (p=0.014) with horses increasing sensitivity over repeated PA measures. The FFS showed no significant differences (p=0.347; ICC 94.7%) in repeated measures with excellent reliability and consistency.

Conclusions: PA showed a lack of consistency in intra-rater reliability conflicting with previous research findings, whereas the FFS showed greater reliability in comparison, however, it proved difficult to use in clinical practice. Manual palpation by physiotherapists was shown to have excellent inter-rater reliability when using a categorical scoring system.

Keywords

Pressure Algometry; Manual Palpation; Horse; Flexiforce Sensor; Objective Measurement

Highlights

- Manual palpation using a categorical scoring system was shown to have excellent inter-rater reliability
- Pressure Algometry lacked consistency in intra-rater reliability
- Flexiforce Sensor showed better consistency for intra-rater reliability then the PA
- 50% of horses showed an increase in sensitivity to the Pressure Algometer
- A categorical scoring system is recommended to aid objectivity in manual palpation

1.0 INTRODUCTION

Epaxial soft tissue pain is suggested to be the greatest global disability of humans [1]. It is also of clinical significance in horses, with incidences reported to be as high as 68% in racehorses [2, 3]. However, diagnosis of equine back pathologies remains challenging [5], due to a wide variety of clinical signs. It has been shown thatup to 74% of lameness cases have associated back pain [4], Clinical findings on diagnostic imaging do not necessarily correlate with the presenting pain and findings observed in horses with no subjective clinical symptoms [6]. However advanced imaging techniques can help to improve objectivity of assessment and diagnosis or rule out pathologies [7, 8]. Imagining techniques such as ultrasound or MRI can be used to assess for soft tissue pathology. During a clinical assessment without the benefit of imaging, manual palpation is an extensively used tool to assess horses and is a major component of training for both veterinarians and physiotherapists [9]. Manual palpation is the ability to feel and assess superficial structures using one or both hands; an essential part of patient assessment and treatment [10]. The ability to correlate pain on manual palpation with soft tissue or bony pathology using more objective and consistent methods would benefit diagnostics and equine pathology management. Yet despite this, current practice is still focused on subjective interchangeable language descriptors and techniques which lack a rigorous evidence base [11].

Pressure Algometry is an established validated tool for the measurement of mechanical nociceptive thresholds (MNT) which relate to the point of biometric pressure application where a behavioural response is noted; the technique has been found to be repeatable in humans and horses [11, 12, 13]. In addition to the pressure algometer (PA), other tools such as the Flexiforce Sensor (FFS) have been used in assessment of pressure [14] and both these methods can be used to measure MNT and potential behavioural responses. The nociceptive threshold is the application of a quantifiable stimulus to a specific area of the body until a physiological or behavioural response is detected, after which the stimulus is then removed [15]. The integration of MNT assessment is a crucial component when examining potential pain and behavioural responses upon palpation testing in animals, as it gives feedback regarding pain sensitivity of an area and how the tissue responds to the application of pressure.

In the human field, pain scoring systems are often used to help the clinician understand patient's pain levels. One example is the Visual Analogue Scale (VAS) scored 0-10 which was established for humans by Hayes and Patterson in 1921[16]. These are scales with results based on the verbal feedback given by the human subject. Horses cannot provide verbal feedback and therefore multiple palpation scoring scales have been established that rely on scoring via an examiner (for examples refer to: [11, 17, 18, 19]. These can be used to score pain, tissue texture and responses, and range from detailed systems which are less open to subjective bias [11, 17] to more basic examples such as Jepsen *et al.* (2006) [20] who used normal, mild, medium and marked for mechanosensitivity on palpation. Anecdotally, detailed scales tend not to be widely used in equine practice, potentially due to time constraints and a lack of awareness [21], which can lead to significant subjective interpretation and variability between clinicians..

Objectivity and consistency within the equine physiotherapy industry can often be unsatisfactory due to the lack of outcome measures, reliable tools and absence of feedback from the animals [11, 22]. There are limited objective outcome measures within the equine industry and only a few studies have investigated the validity and reliability of outcome measures or compared multiple measures. Therefore the aim of this study was to establish if a relationship existed between PA, FFS and manual palpation, and to test intra- and inter-rater reliability of examiners ratings of response to palpation in equine thoracic epaxial region using clinically normal horses.

2.0 METHODS

2.1 Subjects

Twenty two gelded horses of mixed breeding, age and height, based at a university equine centre were used with 10 randomly selected for the inter-rater testing (Group 1) and 22 participating in the intra-rater (Group 2) component of the research. All horses had undergone a full veterinary work up and health check in the last six months prior to the research starting and were all in full work and health monitored as per the yard protocol. Each horse also underwent a clinical assessment and was assessed by observation and palpation for any muscle spasm or significant asymmetries prior to the taking part in the study by a qualified Association of Chartered Physiotherapists in Animal Therapy (ACPAT) Chartered Physiotherapist. Any horse which showed clinical signs of pain, muscle spasm or abnormalities was excluded from the research. Four horses fell into one of the above mentioned exclusion criteria therefore only 22 progressed to data collection.

Mean height for Group 1 was 147.6±1.8cm and group 2, 153±2.8cm. Mean age (years) for group 1 was 14.1 (range, 10-18) and for group 2 12.7 (range, 5-22). Body condition scoring averaged at 3 out of 5 (range 2+-4-) [23]. All horses were tested in the morning in their normal environment prior to any exercise. All horses underwent the same daily management and level of exercise, including a mixture of disciplines on multiple surface types, as per the normal yard routine.

2.2 Data Collection

Three ACPAT Chartered Physiotherapists were trained verbally and practically in the use and application of a PA[1] (Model FDK40) with a $1cm^2$ tip and a Tekscan Flexiforce Sensor with a $1.4cm^2$ moderate pressure sensor (FFS)[2] by a trained research assistant who was experienced in the use of each piece of equipment. Each physiotherapist carried out practice on a hard surface, including three repeats to take a mean for the control, and once on a horse. All three physiotherapists were experienced in assessing equine pain reactions, behaviours and handling horses, with a range of 6 to 20 years' experience. The 14th thoracic dorsal spinous process (T14) was located [12] and the test location was marked using 1cm of breathable, hypoallergenic tape placed at 9cm ventrally from the spinous process of T14, to not interfere with the test site located 10cm ventrally, on the right hand side (Figure 1a) [11]. When using the PA (Figure 1b&1c) and FFS, the dial and the laptop screen were faced away from the experimenters' view to blind the physiotherapists to the data and prevent bias. A research assistant read the readings from the equipment and documented all data. Both pieces of equipment were placed perpendicular to the horse (Figures 1a, 1b & 1c) and a constant pressure was applied for up to two seconds until the MNT was reached and repeated consecutively with between 5-7 seconds between each repeat [11]. If the equipment slipped or the horse moved for reasons unrelated to the test, the measurement was ceased and then repeated.

[Figure 1]

Group 1 horses were tested using manual palpation by the physiotherapists; using the index finger on the dominant hand, applying a moderate pressure for two seconds and asked to evaluate the response, referring to the scoring system (Table 1), disclosing the score to the research assistant by pointing to the chart, to exclude possible researcher bias, by over hearing the score, where the other researchers were present. The order of data collection using the PA, FFS and manual palpation was randomised, and the order of each researcher was also randomised. PA and FFS were used by the three physiotherapists, blinded to their own and others data, and each test was repeated three times and an average taken. Only one repeat was taken for manual palpation in both groups as this was a test for reliability not change and the researchers would have known their score from the first repeat and therefore would have produced bias.

[Table 1]

To examine the intra-rater reliability and the relationship between manual soft tissue palpation, PA and the FFS for group 2, 12 additional horses were tested using the same protocol, by one physiotherapist, to make up a total of 22 subjects. The order of horses and the order of data collection using the PA, FFS and manual palpation was randomised

2.3 Data Analysis

Data for inter-rater and intra-rater reliability were tested for normality with Kolmogorov-Smirnov test against the significance value of $p < 0.05$ and were found to be a combination of parametric and non-parametric data. Therefore a non-parametric distribution was assumed due to its greater statistical strength over parametric tests [24, 25]. Friedman's Anova identified if any differences occurred in the three measures between the three physiotherapists' measurements for group 1 horses and across measurements for the single physiotherapist in group 2 horses (alpha: $P<0.05$). Subsequent post-hoc Wilcoxon Signed Rank analyses *(*Bonferroni adjusted alpha $p<0.02$) established where differences occurred between trials within the second group of horses. Intra-class correlation coefficient (ICC) estimates and their 95% confident intervals were calculated for inter- and intra-reliability of the physiotherapists and measured using Statistical Package for the Social Sciences (version 23)[3] based on a mean-rating ($k = 3$), absolute-agreement, 2-way mixed-effects model.

3.0 RESULTS

3.1 Group 1: Inter-Rater Reliability

No significant differences were found between physiotherapists when assessing the sample (n=10) by manual palpation, as well as excellent inter-rater reliability ($\chi2(2)=0.875$, $p=0.646$; ICC: 0.900; CIs: 0.76-0.973). Significant differences were found between the values recorded by the three physiotherapists for the PA (p =0.002) and the FFS (p=0.025) as shown in table 1.

[Table 2]

3.2 Group 2: Intra-Rater Reliability

A significant difference (*p* = 0.014) occurred between measurements for the single physiotherapist using PA (n=22). Post-hoc analysis identified significant differences between the first and third PA readings (z = -2.333, p = 0.020) and second and third PA readings (z = -2.012, p = 0.044) but no significant differences were found between the first and second readings (z = -1.090, p = 0.276). Eleven of the subjects (50%) showed decreased MNT values (increased sensitivity), demonstrating a behavioural response to less force applied via the PA and 13.64% had increased MNT values (decreased sensitivity) over the three repeats

No significant differences were found (*p* = 0.347) between FFS scores for a single rater. ICC estimates using absolute agreement showed excellent intra-rater reliability [26] ($\chi2(2)$=0.347, ICC: 0.947; CIs: 0.891-0.976).

For the majority of horses tested FFS readings remained constant (77.23%) across repeated trials, of the remaining horses 9.1% (n=2) recorded increased sensitivity to FFS and 13.63% (n=3) (Table 3) recorded a decrease in sensitivity as trials progressed sequentially.

[Table 3]

4.0 DISCUSSION

No significant differences were found between manual palpation scores across the three physiotherapists. This suggests that manual palpation skills are reliable which may be as a result of prior training with the scoring system. Although a categorical scale, the descriptors allowed the physiotherapists to classify the response to the application of moderate pressure. The use of a scoring system by clinicians would therefore be advocated in practice, despite disagreeing with previous literature that suggested large inconsistencies exist between clinicians [12]. Interestingly Seffinger *et al.* (2004) [27] concluded that intra-rater reliability is superior to inter-rater, however results from this research study suggest otherwise. Mechanical equipment is reported to be more repeatable and accurate than manual palpation [28] however our results suggest that within this sample and compared to experienced physiotherapists, the equipment, lacks reliability in comparison to manual palpation. Due to the differences in MNT reported by the three physiotherapists in this study no further analysis of agreement between PA and manual palpation score could be undertaken.

PA and FFS inter-rater data showed a large range, both with significant differences between repeated measures for each physiotherapist, for instance horse 1 recorded both the lowest PA reading (4.8kg/cm^2) and highest (8.2 kg/cm^2) reducing agreement between the measurers. This study used a small sample size due to time constraints on the yard, availability of the horses time as they were needed for work and on limited researcher availability, only 10 horses could be used (n=10) and the significant differences found between physiotherapists using the PA conflict with previous human based literature [29]. However, it should be noted that within these human based trials patient responses to the PA were measured with verbal feedback from the subject. In contrast investigation into equine subjects requires the accurate appraisal of non-verbal communication which can lead to greater subjectivity. However, the risk of false results due to patient bias is therefore limited. The range of PA readings was lower than previously reported by Haussler and Erb (2006a&b) [12, 13] who noted a mean of 10.5 kg/cm^2 in non-painful sites versus 5.5 kg/cm^2 in painful sites and when initially testing the use of the PA on horses; a mean of 12 kg/cm^2. In a study assessing sacro-iliac joint region pain PA readings were also higher (sample mean 5.3kg/cm^2) [11] than the mean in this study (4.1kg/cm^2). The lower MNT in this study may be due to differences in the rate of force applied or agreement on behavioural response that constitutes the MNT, although neither factor has been assessed empirically.

Post-hoc testing showed significant differences between repeated applications of the PA. Fifty percent of subjects showed an increase in sensitivity to PA and mean MNT scores decreased over three repeats. Haussler and Erb (2006b) [13] reported a similar increase in sensitivity (24%). FFS, for the three repeats, showed consistent results with only 9.1% showing increased sensitivity to FFS in relation to MNT. This is a significantly lower percentage when compared to PA results and therefore shows the FFS to be less sensitising with sequential application thus increasing the rationale for its use in assessment of epaxial muscles.

The FFS was shown to have no significant differences in intra-rater reliability in this research study. This could be due to the FFS being more comparable in reality to a physiotherapist’s palpation, allowing for the physiotherapist to feel changes and the MNT quicker or feel more subtle changes than the PA. In addition the FFS is more comparable to human manual palpation with a greater surface area ($1.4cm^2$) when compared to PA ($1cm^2$).

The manual palpation scale used in this research paper was easy to use and follow and was very similar to the Varcoe-Cocks *et al.* (2006) [11] scale. In comparison to previous papers and scales used [19, 20] it was significantly more detailed and therefore could be argued to be less open to subjective interpretation, however further testing between greater numbers of assessors and in horses diagnosed with back pain is required to confirm this. Despite supposed subjectivity [5] with a detailed scoring system and using advanced trained clinicians, scoring by manual palpation appears to be reliable to be reliable, supporting its use in evidence based practice. Future research comparing manual palpation sensitivity when back pain has been diagnosed would now be valuable.

4.1 Limitations

Despite greater reliability than the PA the practicalities of the FFS in a clinical setting were difficult. The FFS involved having a laptop to receive the result through Wi-Fi, therefore was dependent on a signal being achievable. It required several calibrations intermittently throughout data collection and retracting data from the laptop was arduous. The PA used was a manual force gauge identical to that used in other PA studies, however it could lack sensitivity at low numbers and could have an effect on consistency. The gauge on the manual PA system was at times difficult to interpret, which may have affected accuracy of readings. A digital PA or other similar devices would help limit variables and improve accuracy of results.

Longitudinal studies could also be utilised with a number of professionals follow specific manual palpation scales and qualitative feedback given and analysed exploring a collaboration of systems to produce one distinctive system which could be used by veterinarians and physiotherapists as a gold standard. In addition, further research with greater numbers may be argued to help increase the strength of the results and more training should be given the researchers in the use of the equipment to help increase the repeatability between researchers.

5.0 CONCLUSION

This research has added to the debate over the use of the PA as an established measurement tool as it showed the PA lacked consistency in intra-rater reliability . Other objective scoring methods for palpation such as the FFS which may represent a more realistic likeness to manual palpation, being more subtle with regards to the behavioural response. Research with larger sample sizes could further investigate FFS to explore validity and reliability and its use in the equine industry, however such tools need to be practical to use 'in the field'. Manual palpation scoring, using a categorical rating scale, carried out by physiotherapists, was shown to have inter-rater reliability. It is therefore recommended that a simple palpation scoring system should be used in clinical practice with the aim of increasing objectivity during palpation assessment.

Manufacturers Address'

[1]Wagner Instruments Inc., Greenwich, Connecticut, USA

[2]Tekscan Inc., South Boston, MA, USA

[3]SPSS Inc., Chicago, IL, USA

REFERENCES

[1] Hoy D, March L, Brooks P, Woolf A, Blyth F, Vos T, et al. Measuring the global burden of low back pain. Best Practice and Research in Clinical Rheumatology. 2010 24(2):155-165.

[2] Stubbs N, Riggs C, Hodges P, Jeffcott L, Hodgson D, Clayton H, et al. Osseous spinal pathology and epaxial muscle Ultrasonography in Thoroughbred racehorses. Equine Veterinary Journal. 2010 42(3):654-661.

[3] Jeffcott LB. Disorders of the thoracolumbar spine of the horse- a survey of 443 cases. Equine Veterinary Journal. 1980 12:197-210.

[4] Burns G, Dart A, Jeffcott L. Clinical progress in the diagnosis of thoracolumbar problems in horses. Equine Veterinary Education 2018 30(9):477-485

[5] Landman M, De Blaauw JA, Hofland LJ, Van Weeren PR. Field study of the prevalence of lameness in horses with back pain. Veterinary Record. 2004 155(6):165-168.

[6] Nagy A, Dyson S, Barr A. Ultrasonographic findings in the lumbosacral joint of 43 horses with no clinical signs of back pain or hindlimb lameness. Veterinary Radiology and Ultrasound [Online]. 2010 51(5):533-539.

[7] Labruyere J, Schwarz T. CT and MRI in veterinary patients: an update on recent advances. British Medical Journal. 2013 35:546-563.

[8] Puchalski SM. Advances in equine computed tomography and use of contrast media. Veterinary Clinics: Equine Practice. 2012 3:563-581.

[9] McGowan C, Goff L. Animal Physiotherapy: Assessment, Treatment and Rehabilitation of Animals. 2nd ed. West Sussex: John Wiley and Sons Ltd; 2016.

[10] Paulet T, Fryer G. Inter-examiner reliability of palpation for tissue texture abnormality in the thoracic paraspinal region. International Journal of Osteopathic Medicine. 2009 12(3):92-96.

[11] Varcoe-Cocks K, Sagar KN, Jeffcott LB, McGowan CM. Pressure algometry to quantify muscle pain in racehorses with suspected sacroiliac dysfunction. Equine Veterinary Journal. 2006 38(6):558-562.

[12] Haussler KK, Erb HN. (2006a) Pressure algometry for the detection of induced back pain in horses: a preliminary study. Equine Veterinary Journal. 2006a 38(1):76-81.

[13] Haussler KK, Erb HN. Mechanical nociceptive thresholds in the axial skeleton of horses. Equine Veterinary Journal. 2006b 38(1):70-75.

[14] Ferguson-Pell M, Hagisawa S, Bain D. Evaluation of a sensor for low interface pressure applications. Medical engineering & physics. 2000 Nov 1;22(9):657-63.

[15] Love EJ, Murrell J, Whay HR. Thermal and mechanical nociceptive threshold testing in horses: a review. Veterinary Anaesthesia and Analgesia. 2011 38(1):3-14.

[16] Hayes MHS, Patterson DG. Experimental development of the graphic rating method. Psychological Bulletin. 1921 18:98–99.

[17] De Heus P, Van Oossanen G, Van Dierendonck MC, Back W. A pressure algometer is a useful tool to objectively monitor the effect of diagnostic palpation by a physiotherapist in warmblood horses. Journal of Equine Veterinary Science. 2010 30(6):310-321.

[18] Bendtsen L, Jenson R, Jenson NK, Olesen J. Pressure-controlled palpation: a new technique which increases the reliability of manual palpation. Cephalalgia. 1995 15:205-210.

[19] Sakai F, Ebihara S, Akiyama M, Horikawa M. Pericranial muscle hardness in tension-type headache: A non-invasive measurement method and its clinical application. Brain [Online]. 1995 118:523-531.

[20] Jepsen JR, Laursen LH, Hagert CG, Kreiner S, Larsen AL. Diagnostic accuracy of neurological upper limb examination I: inter-rater reproducibility of selected findings and patterns. Manual Therapy. 2006 6:8.

[21] Tabor G, Williams J. The use of outcome measures in equine rehabilitation. The Veterinary Nurse. 2018 *9*(9):497-500.

[22] Ranner W, Gerhard H. Diagnostik bei verdacht auf ruckenerkrankungen beim pferd. Pferdeheilkunde. 2001 3:225-232.

[23] Carroll CL, Huntington PJ. Body condition scoring and weight estimation of horses. Equine Veterinary Journal. 1988 20(1):42-45.

[24] Field A. Discovering Statistics using IBM SPSS statistics. 4th ed. London: SAGE Publications Ltd; 2017.

[25] Whitley E, Ball J. Statistics review 6: nonparametric methods. Critical Care. 2002 6(6):509-513.

[26] Koo TK, Lee MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine. 2016 15(2):155-163.

[27] Seffinger MA, Najm WI, Mishra SL, Adams A, Dickerson VM, Murphy LS, et al. Reliability of spinal palpation for diagnosis of back and neck pain: A systematic review of the literature. Spine. 2004 29(19):413-425.

[28] Gwilliam JC, Pezzementi Z, Jantho E, Okamura AM, Hsiao S. Human vs. Robotic tactile sensing: Detecting lumps in soft tissue. In: Haptics Symposium. USA Haptics Symposium: IEEE; 2010.

[29] Fingleton CP, Dempsey L, Smart K, Doody CM. Intraexaminer and interexaminer reliability of manual palpation and pressure algometry of the lower limb nerves in

asymptomatic subjects. Journal of Manipulative Physiological Therapy. 2014 37:97-104.

Figure 1

	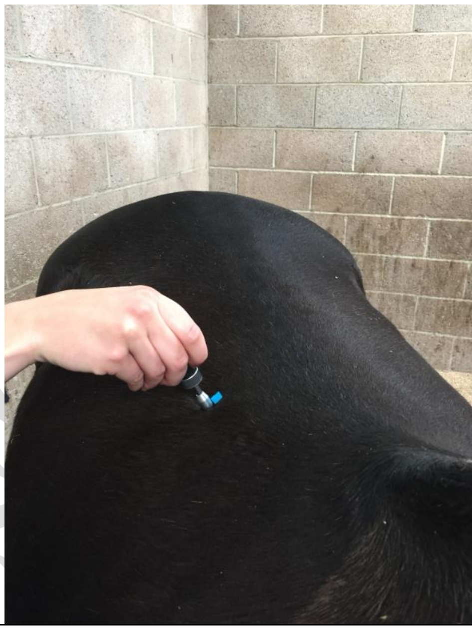
Figure 1a: Location of palpation site	Figure 1b: PA testing The unit was held perpendicular to the subject and pressed to apply pressure

Figure 1c: The pressure was kept on the PA until a MNT response was obtained. The PA was then removed and the reading obtained from the dial taken down by the research assistant. Note the PA dial faced away from the rater.

Figure 1d: FFS in use on a subject. Pressure was applied over the sensor using the index finger on the clinician dominant hand until a MNT response was obtained and was then removed and a reading taken from the laptop by the research assistant.

Figure 1 Use of a Pressure Algometer (PA) (1b & 1c) and a Flexiforce Sensor (FFS) (1d) in the equine epaxial region, 10cm lateral to the dorsal spinous process of the 14th thoracic vertebrae to assess the mechanical nociceptive threshold (MNT).

[Table 1]

Table 1. Scoring scale used in this research, modified from Varcoe-Cocks *et al.* (2006) and the Modified Ashworth Scale (Ravara *et al.* 2015)

Score	Description
0	Soft, low tone
1	Normal
2	Increased muscle tone but not painful
3	Increased muscle tone and/or painful (slight associated spasm on palpation, no associated movement
4	Painful (associated spasm on palpation with associated local movement, i.e. pelvic tilt, extension response),
5	Very painful (spasm plus behavioural response to palpation, i.e. ears flat back, kicking).

[Table 2]

Table 2: Pressure Algometry (PA) (kg), Flexiforce Sensor (FFS) (g) and Manual Palpation scores for inter-rater testing (group 1), by researchers (R) 1, 2 and 3 with repeated measures one (a), two (b) and three (c).

	PA	PA	PA		FFS	FFS	FFS		Manual Palpation
Horse	**R1a**	**R1b**	**R1c**	**Mean**	**R1a**	**R1b**	**R1c**	**Mean**	**R1a**
1	8.2	5.6	5.2	6.3	2566.7	2133.3	2933.3	2544.4	2
2	3.0	3.0	3.2	3.1	1433.3	1466.7	1133.3	1344.4	3
3	3.4	3.1	3.7	3.4	1419.2	2596.2	1488.5	1834.6	3
4	4.8	4.8	4.6	4.7	1834.6	1765.4	2457.7	2019.2	4
5	4.1	3.7	3.9	3.9	2803.8	2492.3	3011.5	2769.2	2
6	6.2	4.4	4.8	5.1	2111.5	2146.2	2734.6	2330.8	3
7	5.5	5.3	4.4	5.1	1592.3	1315.4	1384.6	1430.8	1
8	4.2	4.2	4.0	4.1	1315.4	969.2	900.0	1061.5	3
9	4.2	4.0	3.8	4.0	1488.5	1176.3	1211.5	1292.1	4
10	5	5.2	4.9	5.0	1073.0	934.6	1003.8	1003.8	2
	R2a	**R2b**	**R2c**	**Mean**	**R2a**	**R2b**	**R2c**	**Mean**	**R2**
1	7.0	7.8	7.6	7.5	1666.7	1666.7	2133.3	1822.2	2
2	3.1	3.0	3.0	3.0	1488.5	969.2	1073.0	1176.9	3
3	3.7	3.0	3.0	3.2	2146.2	761.5	1211.5	1373.1	3
4	4.0	4.8	4.2	4.3	1142.3	1038.5	1384.6	1188.5	2
5	3.2	3.8	3.3	3.4	934.6	1315.4	1280.8	1176.9	2
6	3.0	3.7	4.1	3.6	1557.7	1038.5	761.5	1119.2	3
7	5.0	5.2	4.8	5.0	1073.0	830.8	657.7	853.8	2
8	3.0	3.0	3.0	3.0	1038.5	1038.5	1107.7	1061.6	3
9	3.0	3.0	3.0	3.0	519.2	830.8	484.6	611.5	3
10	3.2	3.1	3	3.1	1107.7	726.9	865.4	900.0	2
	R3a	**R3b**	**R3c**	**Mean**	**R3a**	**R3b**	**R3c**	**Mean**	**R3**
1	6.2	6.0	4.8	5.7	3033.3	2633.3	2066.7	2577.8	2
2	3.0	3.0	3.0	3.0	1696.2	1557.7	1523.0	1592.3	3
3	3.0	3.1	3.0	3.0	1384.6	2457.7	1800.0	1800.8	4
4	3.0	3.1	3.8	3.3	2215.4	1765.4	796.2	1592.3	3
5	3.8	3.7	3.2	3.6	2388.5	1488.5	1315.4	1730.8	2
6	5.6	4.4	5.1	5.0	3253.8	2838.5	2215.4	2769.2	1
7	3.0	4.7	4.2	4.0	1765.4	1834.6	1696.2	1765.4	1
8	3.8	3.0	3.0	3.3	1350.0	761.5	588.5	900.0	3
9	3.0	3.0	3.0	3.0	761.5	346.2	588.5	565.4	3
10	4.4	3.8	3.5	3.9	2180.8	1730.8	1038.5	1650.0	2

[Table 3]

Table 3: Pressure Algometry (PA) (kg), Flexiforce Sensor (FFS) (g) and Manual Palpation for intra-rater testing (group 2) by Researcher (R1) with repeated measures one (a), two (b) and three (c).

	PA				FFS				Manual Palpation
Horse	**R1a**	**R1b**	**R1c**	**Mean**	**M1**	**M2**	**M3**	**Mean**	**M**
1	8.2	5.6	5.2	6.33	2566.7	2133.3	2933.3	2544.43	2
2	3.0	3.0	3.2	3.06	1433.3	1466.7	1133.3	1344.43	3
3	3.4	3.1	3.7	3.4	1419.2	2596.2	1488.5	1834.63	3
4	4.8	4.8	4.6	4.73	1834.6	1765.4	2457.7	2019.23	4
5	4.1	3.7	3.9	3.9	2803.8	2492.3	3011.5	2769.2	2
6	6.2	4.4	4.8	5.13	2111.5	2146.2	2734.6	2330.76	3
7	5.5	5.3	4.4	5.06	1592.3	1315.4	1384.6	1430.76	1
8	4.2	4.2	4.0	4.13	1315.4	969.2	900.0	1061.53	3
9	4.2	4.0	3.8	4.0	1488.5	1176.3	1211.5	1292.1	4
10	5.0	5.2	4.9	5.03	1073	934.6	1003.8	1003.8	2
11	6.0	4.6	4.4	5.0	813.8	661.2	762.9	745.97	2
12	2.6	4.6	4.2	3.8	712.0	712.0	813.8	745.93	2
13	6.0	5.4	4.6	5.33	762.9	813.8	915.5	830.73	3
14	4.6	5.0	2.4	4.0	2085.3	2594	2136.2	2271.83	2
15	5.2	4.6	4.4	4.73	1373.3	1220.7	1017.2	1203.73	3
16	2.6	2.4	2.0	2.33	1678.4	1322.4	1576.7	1525.83	2
17	8.0	9.4	8.0	8.47	533.7	440.3	440.3	471.43	1
18	5.2	5.6	6.0	5.6	1576.7	1678.4	1271.6	1508.9	4
19	6.0	5.0	4.6	5.2	1983.6	1678.4	1780.2	1814.06	2
20	4.0	4.2	2.6	3.6	1322.4	966.4	1068.1	1119.0	3
21	5.2	4.0	4.6	4.6	2390.5	2441.4	2237.9	2356.7	3
22	5.0	6.0	5.6	5.53	1322.4	1678.4	1373.3	1458.0	2